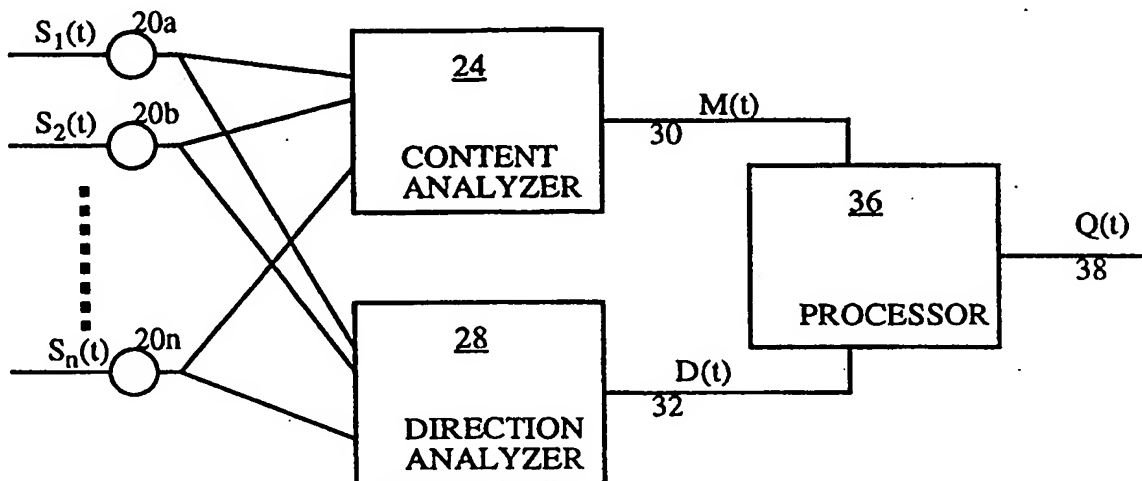




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>G01V 1/00, G01H 3/08, G10L 3/02</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 96/41214</b> <b>(43) International Publication Date:</b> 19 December 1996 (19.12.96)
<b>(21) International Application Number:</b> PCT/US96/09034 <b>(22) International Filing Date:</b> 6 June 1996 (06.06.96)  <b>(30) Priority Data:</b> 08/485,368                      7 June 1995 (07.06.95)                      US  <b>(71) Applicant:</b> SENSIMETRICS CORPORATION [US/US]; 26 Landsdowne Street, Cambridge, MA 02139 (US).  <b>(72) Inventors:</b> DURLACH, Nathaniel, I.; 21 Partridge Hill Road, Weston, MA 02193 (US). BERKOVITZ, Robert, A.; 45 Longwood Avenue, Brookline, MA 02146 (US).  <b>(74) Agents:</b> STEINBERG, Donald, R. et al.; Hale and Dorr, 60 State Street, Boston, MA 02109 (US).		<b>(81) Designated States:</b> CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>

(54) Title: APPARATUS AND METHOD FOR SPEECH RECOGNITION USING SPATIAL INFORMATION



## (57) Abstract

A speech recognition system combines content analysis of a signal with directional analysis of the signal to calculate probable contents of the signal and the probability that the contents are as calculated. Content analysis can be combined with independent source direction analysis, consisting of the probable direction of the source and the probability that the source is as determined, to refine the content analysis. The direction analysis can affect either the calculation of the probable contents of the signal or the calculation of the probability that the contents are as calculated. The improved information obtained by combining the content analysis with directional analysis can be used to improve subsequent content analyses or directional analyses.

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

## APPARATUS AND METHOD FOR SPEECH RECOGNITION USING SPATIAL INFORMATION

### Field of the Invention

5           This invention relates to the recognition of intelligent information in signals. More particularly, it relates to speech recognition in environments that contain interference.

### Background of the Invention

10           Automatic speech recognition systems typically perform poorly when operating in noisy environments, where they must distinguish between speech ("target") signals and noise (interference). If the acceptance criteria are adjusted so that a high proportion of target signals are accepted, then many of the interfering noise signals are also accepted. Consequently, the false-alarm rate is high. On the other hand, if the acceptance criteria are adjusted to achieve a low false-alarm rate, then many target signals are rejected. Consequently, the target acceptance  
15           rate is low. Thus, no matter how the acceptance criteria are selected, performance is degraded by the presence of the interference.

          Prior art attempts to solve this problem have been based on increasing the ratio of target signal power to interference signal power. This is accomplished typically by use of one or more of the following three methods. First, input signals  
20           can be filtered to attenuate frequencies not contained in the spectra of the target signals. Although generally helpful, this method is unsatisfactory in noisy environments.

          Second, the microphones can be located close to the source of the target signals to increase the relative strength of these signals. This method is useful only  
25           under limited conditions. It is not effective if the target signal moves or if there are a number of target sources at different locations. Even for a single, fixed source, this method may be unacceptable. For example, the target source may be a person who does not wish to wear a microphone headset. Or, it may be desired that the target not be aware of the signal acquisition system.  
30           

          A third method is to employ a microphone system that is highly directional, oriented so that it is most sensitive in the direction of the target signals. However,

this method requires multiple microphones, placed at different locations. This requires the use of a substantial amount of space, particularly if the system is to be highly directional at low frequencies. Also, movement of the target source in such a system will degenerate performance of the speech recognition system because  
5 the directional pattern of the microphones causes the received signal to vary with respect to the position of the target. As a result, movement of the target source causes changes in the received signal that cannot be distinguished from changes in the target signal. This distortion of the received signal occurs even without any noise.

10 It is therefore an object of the present invention to provide an improved speech recognition system.

It is another object of the present invention to provide a speech recognition system with improved performance in environments that contain interference.

It is a further object of the present invention to provide a speech recognition system with improved performance in environments where the target source  
15 moves.

#### Summary of the Invention

According to the present invention, these and other objects and advantages  
20 are achieved by employing a diversified microphone system to obtain information on the direction of the target sound source (a "directional signature"). This directional information is combined with the content information on which automatic speech recognition is customarily based.

The system employs two subsystems. The first includes a microphone system and associated signal processing to determine the direction (in terms of the angular coordinates,  $\theta$  and  $\Phi$ ) of the sound source. The second combines this directional information with the content information to determine whether the received signal should be accepted as a target signal or rejected as interference. The directional information provided by the first subsystem can be used to exclude  
25 interfering signals arising from non-target directions with a low probability that desired target signals will be excluded. The directional information also permits  
30 signals for which the content is less certain to be accepted as long as they are very

likely to come from a target source.

#### Brief Description of the Drawings

Figure 1 is a block diagram of a first embodiment of the speech recognition system of the present invention.

Figure 2 is a block diagram of a second embodiment of the speech recognition system of the present invention.

#### Detailed Description of the Invention

The present invention is for an apparatus and method for speech recognition using spatial information to reduce the effects of interference.

Referring to Figure 1, time-varying signals  $S_1(t) - S_n(t)$  are input to microphones 20a - 20n. Preferably, microphones 20a-20n are omnidirectional microphones or have a wide acceptance angle. These signals are then transmitted to content analyzer 24 and source direction analyzer 28.

Content analyzer 24 is a conventional speech recognition subsystem, which can include dedicated hardware or a specially programmed computer, as is well known in the art. For example, the "Dragon Dictate" system by Dragon Systems of Newton, Massachusetts is suitable for the systems of the present invention.

Content analyzer 24 provides an output  $M(t)$ , on line 30.  $M(t)$  is the estimated most probable message, along with a measure of the associated probability or confidence rating.  $M(t)$  may represent the utterances judged to be the most likely to have been transmitted out of the set of utterances in the system's vocabulary. In a preferred embodiment,  $M(t)$  at each time interval consists of a series of utterances and probabilities  $[(M_1, m_1), (M_2, m_2), \dots (M_k, m_k)]$  in order of decreasing probability or confidence, where the  $M_i$  are the utterances and the  $m_i$  are the corresponding probabilities, where  $m_1 \geq m_2 \geq \dots \geq m_k$  and  $0 \leq m_i \leq 1$  for all  $i$ .

Source direction analyzer 28 provides an output  $D(t)$  on line 32. Where there is assumed to be only one source target signal,  $D(t)$  is the estimated most probable source direction  $(\theta, \Phi)$ , along with a measure of the associated probability or confidence. In a preferred embodiment,  $D(t)$  at each time interval consists of a series of source directions and probabilities  $[(D_1, d_1), (D_2, d_2), \dots (D_k, d_k)]$  in

order of decreasing probability or confidence, where the  $D_i$  are the source directions and the  $d_i$  are the corresponding probabilities, where  $d_1 \geq d_2 \geq \dots \geq d_k$  and  $0 \leq d_i \leq 1$  for all  $i$ . Where there may be multiple target sources, each item in the series could include multiple source locations.

- 5           The probability or confidence level associated with the accuracy of the identification of an utterance (the  $M(t)$  function) is normally derived from an evaluation of the dissimilarity of two feature vectors. This technique is well known in the art and its application is described in the book *Fundamentals of Speech Recognition*, by L. Rabiner and B.H. Juang (Prentice-Hall, 1993).
- 10           The probability measure associated with the source direction (the  $D(t)$  function) can be derived from a histogram record of directions associated with the source for previous utterances recognized by the speech recognition system. For example, the azimuth ( $\theta$ ) associated with each arriving utterance is determined, and a record is maintained of the azimuth associated with each of the last 40 utter-
- 15           ances meeting minimum probabilities. The distribution of the azimuths forms a probability density function. Then, for example, if the measured direction of the source for the most recent utterance has been associated with 25% of the 40 utterances for which records are maintained, a probability or confidence level of 0.25 would be assigned to that azimuth value.
- 20           The histogram -- that is, the record of measured directions and counts -- is continually updated as new utterances are accepted by the system and the data for the oldest utterances is dropped. Consequently, the histogram changes in a manner that adapts to changes in the position of the source. For example, if a number of recent utterances arise from azimuth angles to the right of the centroid of the probability density function, which would correspond to movement of the source to the
- 25           right, the adaptive process causes the centroid to shift to the right.
- The probability determination can be modified by adding a weighting function based on the age of each of the records (*e.g.*, give more weight to the most recent records) or some other factor appropriate to known characteristics of the
- 30           source. Preferably, the weighting function is derived from prior knowledge of the probable locations and movements of the source. This prior knowledge can be based, for example, on physical constraints such as the location of the speaker's

chair, or on information obtained during the phonetic training period normally required by a speech recognition system so that it can take into account the individual characteristics of the speaker.

The source direction where there is only one source location active at any one time can be determined for one angular dimension  $\theta$  by using two microphones, which have inputs  $S_1$  and  $S_2$ . In a preferred embodiment,  $S_1(t)$  will be found to be attenuated and delayed with respect to  $S_2(t)$ , or vice versa. Values of  $\alpha$  and  $\tau$  are determined to minimize the difference signal,  $Z(t) = S_2(t) - \alpha S_1(t-\tau)$ . In the absence of noise, values for  $\alpha$  and  $\tau$  can be selected for which  $Z(t)$  will be zero for all times  $t$ .

This can be accomplished, for example, with an array of delay lines, where each delay line represents a different delay ( $\tau$ ) to the signal  $S_1(t)$ , and an array of attenuation lines, where each attenuation line represents a different attenuation ( $\alpha$ ) of the same signal,  $S_1(t)$ . A matrix is formed, consisting of each combination of a delay and an attenuation of the signal  $S_1(t)$  (i.e.,  $\alpha S_1(t-\tau)$ , for different values of  $\alpha$  and  $\tau$ ). Each combination is compared to the other signal,  $S_2(t)$ . The comparison that yields the smallest difference between the signals (i.e., the minimum value for  $Z(t)$ ) determines the appropriate values for  $\alpha$  and  $\tau$ . The conditions under which the system will be used will determine the appropriate number of delay and attenuation lines needed to obtain sufficient accuracy in the determination of  $\alpha$  and  $\tau$ .

The direction of the source is then determined by converting  $(\alpha, \tau)$  to  $\theta$ , as is well known in the art. This method can be applied where  $n$  source locations may be active at a given time by employing  $n+1$  microphones and determining values for  $\alpha_1-\alpha_n$  and  $\tau_1-\tau_n$  to minimize  $Z(t)$  for all times  $t$ .

Alternatively, the source direction can be determined without the use of delay lines by measuring the delay  $\tau$  between signal onsets in the two channels and using a comparator to determine the difference in amplitude of the signal onsets; by using microphones with different directional patterns and measuring the amplitude difference in the two channels; by measuring the correlation coefficient between the two channels after first whitening the spectra of the signals; or by other methods known in the art. These methods also can be applied where more than one source location may be active at a given time.

The outputs  $M(t)$  on line 30 and  $D(t)$  on line 32 are input to and processed in processor 36. The signals  $M(t)$  consist of pairs  $[(M_1, m_1), (M_2, m_2), \dots (M_k, m_k)]$  for each time segment. The probability  $m$  that the received signal arises from one of the utterances  $M_i$  can be represented by:

5

$$m = \sum_{i=1}^k m_i$$

and the probability that the received signal does not arise from one of the utterances  $M_i$  is given by  $m^* = 1-m$ .

10

Where there is a single signal source, the directional information  $D(t)$  can be represented by the total probability  $d$  that the source is a target source. That is, having determined the direction, the source direction analyzer 28 provides a probability  $d$  that the source is a target. The probability  $d$  is independent of the content information  $M$ . In this situation, the output  $Q(t)$  of processor 36, on line 38, is the

15

set of pairs  $[(M_1, q_1), (M_2, q_2), \dots (M_k, q_k)]$  for each time segment, where  $q_i = m_i d$ . Thus, by combining the directional information  $D(t)$  with the content information  $M(t)$ , the probability of the signal resulting from target message  $M_i$  is  $m_i d$  instead of  $m_i$ . The probability of the signal resulting from some target message is:

20

$$m' = d \sum_{i=1}^k m_i = dm = d(1-m^*)$$

and the probability of the signal not resulting from some target message is given by  $m'^* = 1-dm$ .

25

As shown, the addition of the directional information  $D(t)$  decreases the probability that any signal will be regarded as a target message because probabilities based on both content and direction are considered. However, this effect can be offset by lowering the numerical threshold of acceptance of an utterance, a quantity that is or can readily be made accessible to the user in the software of speech recognition systems.

30

Numerous methods may be employed to determine the probability that the source is a target source from the directional information  $D(t)$ . In a first method,  $a$



*priori* information on the position of the target source is utilized. In one embodiment, the system would assume that the mean target position is straight ahead ( $\theta=0$ ), so that  $d=1$  for  $\theta=0$ , and that the probability of the target being at any other angle decreases exponentially with angle (that is,  $d = e^{-g\theta}$ , where  $g$  is a constant).

5           In a second method,  $d(\theta)$  is constructed during training of the recognition system. For example, where the system is trained in a quiet environment, a histogram of target values  $\theta$  is accumulated and used to determine the function  $d(\theta)$ . In a preferred embodiment using this method, the distribution of  $\theta$  for interference also is estimated during training and included in the function  $d(\theta)$ .

10           In a third method,  $d(\theta)$  is constructed by exploiting the message content information  $M(t)$  during real-time use of the system. For example, for each received signal, the source direction angle  $\theta$  and probability  $m$  that the received signal arises from one of the utterances is used to develop a histogram for acceptable source direction angles  $\theta$ . The histogram is a representation of a list of  
15           ordered pairs  $(\theta_i, m_i)$  in which each pair consists of an angle and a probability, and in which totals have been sorted by  $\theta$ . Since the histogram is derived from actual data, it can be used to replace a gaussian or other nominal statistical distribution.

            The histogram is used to form  $d(\theta)$ , which processor 36 in turn uses to obtain the output  $Q(t)$ . Preferably, this modified content information  $Q(t)$  is fed  
20           back to source direction analyzer 28 to improve the directional information. For example, the modified content information  $Q(t)$  can be used to modify the directional probability function,  $d(\theta)$ . This method can be used where initial values for  $d(\theta)$  are formed from a different method. The modified content information  $Q(t)$  also can be fed back to content analyzer 24 to improve the initial message content  
25           information  $M(t)$ . For example, the modified content information  $Q(t)$  can be used to modify the utterance probability function.

            In addition to improving the message content information  $M(t)$ , the directional information can be used to permit the voice recognition system to distinguish between sources as long as the sources generally do not overlap. For  
30           example, where there are two sources that generally do not overlap, there would be two direction functions  $d_1(\theta)$  and  $d_2(\theta)$ , for which there generally are no angles  $\theta$  for which both  $d_1(\theta)$  and  $d_2(\theta)$  are large. With these functions, the system could

determine which source supplied an utterance. Or, the determination of the most probable message could depend on the direction.

Alternatively, the directional information can be integrated with the content information in other ways. In one embodiment, the estimated direction  $\theta$  is coded  
5 into a signal  $N(t)$ , which is appended to the signals  $S_i(t)$  at the input to content analyzer 24. Preferably, directional segment  $N(t)$  is attached as soon as the directional information is available, as an interruption or a change in the utterance waveform. This transforms the directional information into content information, and the signal is processed without the need for processor 36 to combine directional information with the processed content information. That is, all of the content processing  
10 takes place in content analyzer 24.

In one embodiment, direction signal  $N(t)$  is appended as a suffix to the signals  $S_i(t)$ . Whenever the angle  $\theta$  is outside a previously defined range of acceptable angles, the value of  $N(t)$  is such that content analyzer 24 will reject the signal  
15 as not corresponding to a recognized utterance within the system's vocabulary.

An alternative embodiment for integrating the directional information with the content information is shown in Figure 2. Time-varying signals  $S_1(t) - S_n(t)$  are input to microphones 50a - 50n. These signals are then transmitted to direction analyzer 54 and mixer 56. Direction analyzer 54 provides an output on line 58, in  
20 the manner described above, that reflects the probability that the source signals originate with a desired target source. Line 58 provides the input to noise generator 60, which generates an amount of noise inversely proportional to the probability that the source signals originate with the desired target source. That is, the noise level increases as the probability that the signal arises from the desired target source decreases.  
25

The noise signal is output on line 62, which connects to mixer 56. In mixer 56, the signals  $S_1(t) - S_n(t)$  are mixed with the noise signal from noise generator 60, and then input to content analyzer 64. Content analyzer 64 outputs message information  $Q(t)$  on line 66. An increase in the noise on the signal input to content  
30 analyzer 64 decreases the likelihood that the utterance will be recognized and/or the confidence level in the accuracy of the utterance determination. Thus, represented as noise, the direction information imposes a constraint on the content ana-

lyzing portion of the system.

In a simplified version of the embodiment shown in Figure 2, the direction information would be used to either include or exclude the source signals  $S_1(t)$  -  $S_n(t)$ . In this embodiment, the directional information is used independently of the content information. If the probability that the signal originated with a desired target fell below a predetermined threshold, the system would reject the signal. If the probability exceeded the threshold, the signal would be input to the content analyzer unchanged.

While there have been shown and described examples of the present invention, it will be readily apparent to those skilled in the art that various changes and modifications may be made therein without departing from the scope of the invention as defined by the appended claims. Accordingly, the invention is limited only by the following claims and equivalents thereto.

What is claimed is:

Claims:

1. A recognition system for a source signal comprising:  
input means for obtaining at least one input signal from the source signal;  
a content analyzer having a content input coupled to the input means, a content output, and a content processor for calculating and providing to the content output a content signal representative of an initial calculated probable value of the source signal;  
a direction analyzer having a direction input coupled to the input means, a direction output, and a direction processor for calculating and providing to the direction output a direction signal representative of a calculated source direction of the source signal; and  
a message processor having a first input coupled to the content output, a second input coupled to the direction output, a message output, and a message processor for calculating and providing to the message output a message signal representative of a supplemental calculated probable value of the source signal.
2. The recognition system as in claim 1, wherein the content processor further calculates and provides to the content output a content probability signal representative of a confidence level in the initial calculated probable value.
3. The recognition system as in claim 2, wherein the direction processor further calculates and provides to the direction output a direction probability signal representative of a confidence level in the calculated source direction.
4. The recognition system as in claim 3, wherein the message processor further calculates and provides to the message output a message probability signal representative of a confidence level in the supplemental calculated probable value of the source signal.
5. The recognition system as in claim 4, wherein the input means includes at least one microphone.
6. The recognition system as in claim 4, wherein the message processor calculates the message probability signal by multiplying a value of the direction probability signal by a value of the content probability signal.
7. The recognition system as in claim 4, wherein the message signal is substantially the same as the content signal.

8. The recognition system as in claim 4, wherein the direction processor uses prior calculations of the source direction to calculate the direction probability signal.

9. The recognition system as in claim 4, wherein the direction analyzer further includes a feedback input connected to the message output, and wherein the direction processor uses prior values of the message signal and the message probability signal to calculate the direction probability signal.

10. The recognition system as in claim 4, wherein the content analyzer further includes a feedback input connected to the message output, and wherein the content processor uses prior values of the message signal and the message probability signal to calculate the content probability signal.

11. The recognition system as in claim 4, wherein the message processor calculates the message signal from the content signal and the direction signal.

12. The recognition system as in claim 4, wherein the message processor calculates the message signal from the content signal, the direction signal, and the direction probability signal.

13. A recognition system for a source signal comprising:  
input means for obtaining at least one input signal for the source signal;  
a direction analyzer having a direction input coupled to the input means, a direction output, and a direction processor for calculating a probability that the source signal originates from one of at least one predetermined direction and providing to the direction output a direction probability signal representative of the probability;

a noise generator having an input coupled to the direction output and a noise output, for providing a noise signal inversely proportional to the probability;

a mixer having a first mixer input coupled to the input means, a second mixer input coupled to the noise output, and a mixer output, for combining signals on the first mixer input and the second mixer input; and

a content analyzer having a content input coupled to the mixer output, a content output, and a content processor for calculating and providing to the content output a content signal representative of a calculated value of the source signal.

14. The recognition system as in claim 13, wherein the content processor

further calculates and provides to the content output a content probability signal representative of a confidence level in the calculated value of the source signal.

15. The recognition system as in claim 14, wherein the input means includes at least one microphone.

16. A method for converting a source signal to a symbolic signal comprising the steps of:

obtaining at least one input signal from the source signal;

calculating from the at least one input signal a probable source direction of the source signal;

generating from the calculated probable source direction a direction signal;

appending the direction signal to the input signal;

accepting the appended input signal if the appended input signal corresponds within a predetermined degree to one of a plurality of predetermined acceptable signals;

rejecting the appended input signal if the appended input signal does not correspond within the predetermined degree to one of the plurality of predetermined acceptable signals; and

generating from an accepted appended input signal a signal symbolic of a probable value of the source signal.

17. The method for converting a source signal as in claim 16, wherein the direction signal generating step includes generating a first direction signal if the calculated probable source direction is within a predetermined range of acceptable directions and generating a second direction signal if the calculated probable source direction is not within the predetermined range.

18. The method for converting a source signal as in claim 17, wherein the rejecting step includes rejecting an appended input signal if the appended input signal includes the second direction signal.

19. A method for converting a source signal to a symbolic signal comprising the steps of:

obtaining at least one input signal from the source signal;

calculating from the at least one input signal a probable source direction of the source signal;

generating from the calculated probable source direction a direction signal;  
accepting the input signal if the input signal corresponds within a predetermined degree to one of a plurality of predetermined acceptable signals and the direction signal is within a predetermined range of acceptable direction signals; and  
generating from an accepted input signal a signal symbolic of a probable value of the source signal.

20. A method for converting a source signal to a symbolic signal comprising the steps of:

obtaining at least one input signal from the source signal;  
generating from the at least one input signal a content signal representative of a calculated probable value of the source signal;  
generating from the at least one input signal a direction signal representative of a calculated probable direction of the source signal; and  
generating from the content signal and the direction signal a message signal including a message content component symbolic of a calculated probable value of the source signal.

21. The method for converting a source signal as in claim 20, wherein the content signal generating step includes generating a first content component symbolic of a probable value of the source signal and a second content component representative of a confidence level in the probable value.

22. The method for converting a source signal as in claim 21, wherein the direction signal generating step includes generating a first direction component representative of the probable source direction and a second direction component representative of the confidence level in the probable source direction.

23. The method for converting a source signal as in claim 22, wherein the message signal generating step further includes generating from the second content component and the second direction component a message probability component representative of a confidence level in the message content component.

24. The method for converting a source signal as in claim 22, wherein the message signal generating step includes multiplying a value of the second content component by a value of the second direction component.

25. A method for converting a source signal to a symbolic signal com-

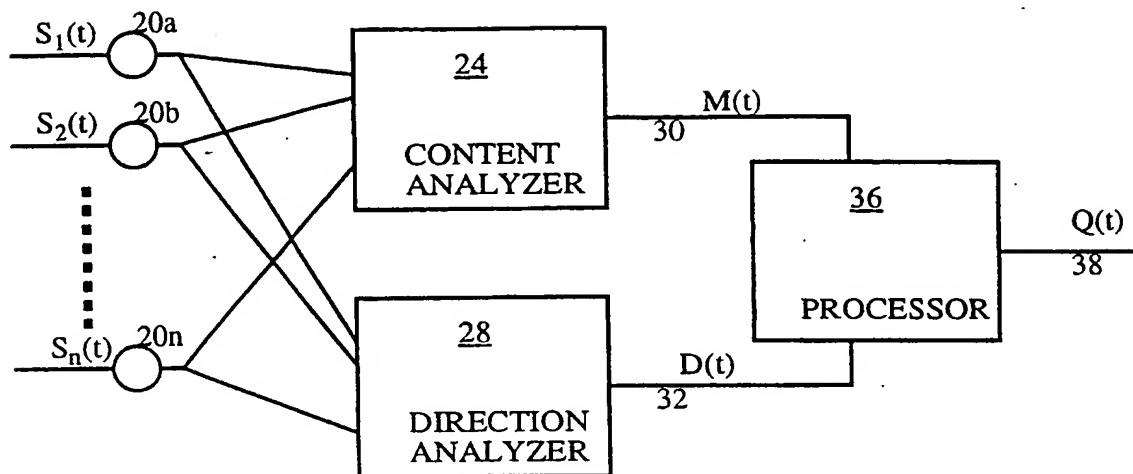
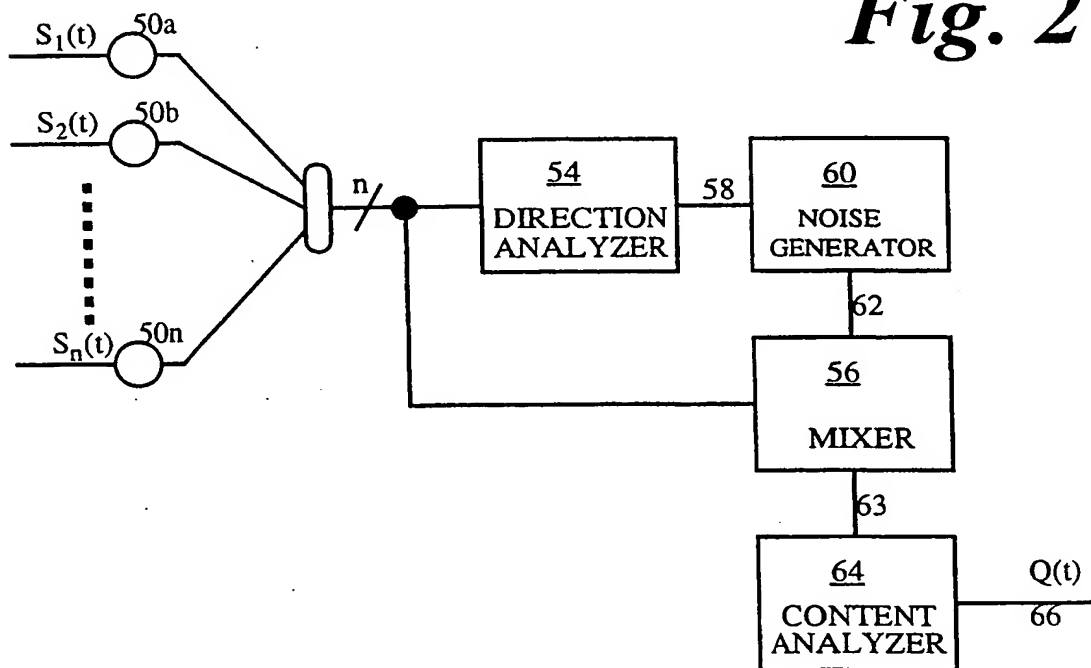
prising the steps of:

- obtaining at least one input signal from the source signal;
- calculating a probability that the source signal originates from one of at least one predetermined directions;
- generating a noise signal inversely proportional to the calculated probability;
- mixing the at least one input signal with the generated noise signal; and
- generating from the mixed signal a content symbol representative of a probable value of the source signal.

26. A method for converting a source signal to a symbolic signal comprising the steps of:

- obtaining at least one input signal from the source signal;
- generating from the at least one input signal a direction signal representative of a calculated probable direction of the source signal; and
- generating from the at least one input signal and the direction signal a content signal representative of a calculated probable value of the source signal.



*Fig. 1**Fig. 2*

# INTERNATIONAL SEARCH REPORT

Intern    nal Application No  
PCT/US 96/09034

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6    G01V1/00    G01H3/08    G10L3/02

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6    G01V    G01H    G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PROCEEDINGS OF THE ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS, PACIFIC GROVE, NOV. 4 - 6, 1991, vol. 1 OF 2, 4 November 1991, CHEN R R, pages 339-343, XP000312748 SWITZER T ET AL: "A CUSTOMIZED BEAMFORMER SYSTEM FOR ACQUISITION OF SPEECH SIGNALS" see page 339, column 1, line 1 - column 2, line 17; figure 1	1,16
Y	EP,A,0 608 708 (IBM) 3 August 1994 see abstract --- -/--	1,16

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- \*&\* document member of the same patent family

Date of the actual completion of the international search

26 September 1996

Date of mailing of the international search report

- 9. 10. 96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+ 31-70) 340-3016

Authorized officer

Haasbroek, J

# INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 96/09034

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, vol. 44, no. 3, 1 June 1995, pages 733-738, XP000527552 HUANG J ET AL: "A BIOMIMETIC SYSTEM FOR LOCALIZATION AND SEPARATION OF MULTIPLE SOUND SOURCES" see abstract; figures 1,2 see page 734, column 1, line 4 - line 16 ---	1,16
A	DE,A,41 41 562 (ATLAS ELEKTRONIK GMBH) 24 June 1993 see column 2, line 1 - line 32 ---	1
A	PATENT ABSTRACTS OF JAPAN vol. 007, no. 212 (P-224), 20 September 1983 & JP,A,58 106667 (FUJITSU KK), 25 June 1983, see abstract -----	1

# INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/US 96/09034

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0608708	03-08-94	US-A- 5491758	13-02-96
		JP-A- 6243297	02-09-94
		US-A- 5544261	06-08-96
		US-A- 5550931	27-08-96
		US-A- 5539839	23-07-96
		US-A- 5544264	06-08-96
DE-A-4141562	24-06-93	NONE	

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**